Stability Study on the Stainless Steel States Kilogram Artifacts Issued by the National Bureau of Standards

Keith R. Arkenberg

Kansas Department of Agriculture,
Division of Laboratories,
Kansas Metrology Laboratory

1. Introduction

The mass scale of the United States has been traditionally determined from the two national kilogram prototypes, K20 and its control standard K4. Periodically these standards are recalibrated at *Bureau of International des Poids et Mesures* (BIPM) as previously described in (Kubarych & Abbott, 2014). From K20 and K4 the United States national kilogram working standards are determined which in turn are used to calibrate the state level kilogram standards (Kubarych, 2000).

In 1965 it was advised that National Bureau of Standards (NBS) issue stainless steel mass artifacts, from 30 kg to 1 mg, to ten state laboratories (Stabler, 1966). The material deemed Checkwate 8 was chosen for this purpose. The austenitic stainless steel round stock was produced by Alleghany Ludlum Steel Corporation using the consumable electrode vacuum melting method with nominal mass-percent composition of: 0.10 carbon, 1.5 manganese, 0.5 silicon, 20.0 chromium, 25.0 nickel, 2.25 molybdenum, 0.2 copper, with the remaining balance of iron resulting in a density near 8.00 g/cm^3 ± 0.05 g/cm^3 at 20 ˚C (Allegheny Ludlum Steel Corporation, 1961). The composition of the standards has proven to be the only one, until recent times, not to exhibit any problems with magnetism (Harris, 2016). Other properties of the Checkwate 8 stock included: corrosion resistance to chemical attack and oxidation, magnetic permeability of 1.007 at 200 H and consistently below 1.01, a cubical coefficient of thermal expansion of 0.000045 per ˚C for temperatures between 20 ˚C and 100 ˚C, and a Rockwell B hardness of Rb 80 (Allegheny Ludlum Steel Corporation, 1961). This round stock was fabricated during three production runs with densities near 8.0015 g/cm^3, 7.9925 g/cm^3, and 7.974 g/cm^3 respectively. From the stock, three companies and finally NBS were chosen to machine the standards: W. &L.E Gurley, Transmetric, Henry Troemner Inc., and NBS Instruments Shop. These standards were produced until the late-1970s and, given the remaining bar stock is no longer in the NBS inventories, now the National Institute of Standards and Technology (NIST), cannot be reproduced or replaced making it impossible to place a modern day value on the sets. At the time of issuance the machining and calibration of each set was near $4500.00 (Keysar, 1973).

2. Method

Stability of the artifacts can be examined from analyzing the historic calibration values issued by NBS and now National Institute of Standards and Technology (NIST). For simplicity and practicality only the kilograms were examined, as these are the reference standards used for mass dissemination and the only artifacts many labs working at the highest level send to NIST for periodic calibration.

Eight state level metrology laboratories were petitioned for all historical calibration values on their NBS issued kilograms and six responded with relevant data sets useful in the study of stability of the artifacts. For the purpose of this paper the true mass values, theoretical mass in vacuum, were examined as reported directly from NBS / NIST calibration reports. The effects of cleaning, determination of density at the time of calibration, and other events cannot be fully accounted for due to the sparse records of such events and will be further examined in Section 3. In Sections 6 and 7 these values are adjusted in an attempt to filter out shifts in the mass scale as maintained by NBS / NIST. By applying these values as corrections the stability of the artifacts is examined without the influence of the stability in the mass scale. For the purpose of this paper the 'stability of the mass scale' will be used to refer to the natural fluctuation in the mass scale due to periodic recalibration of United States kilogram prototype K20.

Each pair of kilograms were assumed to be stored in consistent conditions and used at equal rates to one another. Since it is not known to be case for all of the pairs examined it has to be inferred from the trends seen in the calibration data. Most of the sets exhibit a loss in mass over time due to routine use as would be expected. This excludes calibrations performed between 2010 and 2015 where an upward shift is observed in most of the data sets examined.

To combine the data from the Checkwate 8 kilograms into a single plot some care and considerations had to be considered: how to keep the resolution of the data, how to unify a starting point of the data, and how to obtain a meaningful conclusion from the data. An arbitrary reference value of zero was selected with all subsequent corrections and values offset from the initial reference point. From there, each reported value was determined as a deviation from the initial calibration. Thus the resolution of the data is preserved and to simplify the vertical axis, values are reported in milligram. By examining the deviation one can begin to draw conclusions as to the stability of each pair of artifacts and that of each production run. Finally two methods of reducing the noise in the data are discussed. The first by evaluating the effects of incorporating any reported shifts of calibrations performed in a time period into the data by applying them as corrections to the data and a second by viewing K20 as a perfectly stable artifact with any reported changes in mass due to recalibration summed cumulatively over time and applied as a varying correction over time.

To determine the overall stability of the Checkwate 8 standards each artifact's dataset was examined individually and then as a group. By determining the change in mass between calibration cycles and dividing it by the years between the subsequent calibrations, a change in mass per year was determined for each calibration cycle for each kilogram. Then for each artifact the average change in mass per year for each calibration cycle was averaged to determine the stability of the artifact. Finally each artifact's average stability was averaged resulting in the stability of the Checkwate 8 mass standards, see Table 1.

Figure 1: Long Term Stability of Stainless Steel 1 kg State Standards With Initial Calibration Value Set to Zero

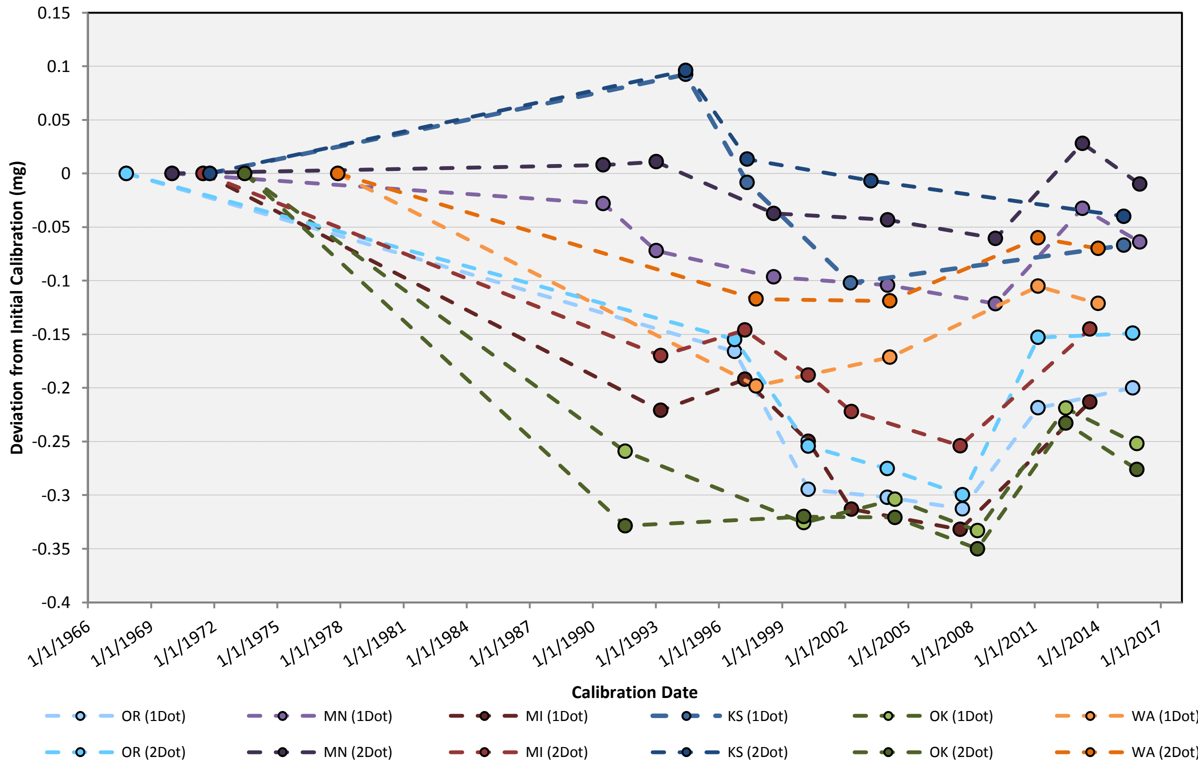

Table 1: Change in Mass per Year

Data Set	Manufacturer and Issuance Year	Density (g/cm^3)	Average Δm/year (mg)	Standard Deviation of Data (mg)
OR (1Dot)	W.&L.E. Gurley	8.0017	0.013	0.015
OR (2Dot)	1967	8.0018	0.015	0.016
MN (1Dot)	Henry Troemner Inc.	7.9926	0.0087	0.0082
MN (2Dot)	1970	7.9926	0.0072	0.0080
MI (1Dot)	Henry Troemner Inc.	7.9925	0.016	0.015
MI (2Dot)	1971	7.9925	0.012	0.006
KS (1Dot)	Henry Troemner Inc.	7.9921	0.015	0.015
KS (2Dot)	1971	7.9925	0.010	0.012
OK (1Dot)	Transmetric	7.9920	0.012	0.008
OK (2Dot)	1973	7.9920	0.011	0.011
WA (1Dot)	NBS Instrument Shops	7.9740	0.0073	0.0028
WA (2Dot)	1977	7.9740	0.0045	0.0034
Maximum Instability	**Henry Troemner Inc.**	**7.9925**	**0.015**	**0.015**
Minimum Instability	**NBS Instrument Shops**	**7.9740**	**0.0045**	**0.0034**
Average Instability	**Checkwate 8**	**7.9740 to 8.0017**	**0.0109**	**0.0036**

The overall stability was determined by averaging the average change in mass per year of each artifact, the standard deviation was determined utilizing the same data.

3. Unknown factors for consideration

Many additional factors need to be considered when examining the stability of the Checkwate 8 artifacts even if they cannot be quantified: surface finish, density determination, cleaning, large changes in mass, handling, and if used on an automated balance.

Surface finish is known to have an effect on the stability of a mass artifact. All sets issued to the state calibration laboratories were refined to a mirror surface. What a mirror surface means is up to the manufacture of the artifact and since surface finish was not and has not been measured on any of the artifacts this will have to be considered consistent.

Density determination is known to affect the stability of a mass artifact due to the cleaning that occurs in the process. It is not known how the artifacts were cleaned. It is known that both the nettoyage-lavage and ultrasonic bath in ethanol methods of cleaning produces very stable results after cleaning and equilibration (Grgic, Grun, Marti, & Berry, 2015). However, if the standard was not allow to equilibrate for the recommended 7 to 10 days after subsequent cleaning the effect could be seen in the reported mass values (Davidson, The use of x-ray photoelectron spectroscopy to estimate the stability of primary mass standards, 2004). Since each pair of kilograms had the density determined prior to the initial calibration and no records exist of the equilibrium process this could be considered as possible explanation for an increase in mass from the first to second calibration, but cannot be accurately accounted for. Most artifacts only had this performed once in their lives except for the standards in Minnesota where the density was determined during the initial calibration and the three concurrent calibrations.

Cleaning that occurred in the state laboratories. From the time of issuance of the Checkwate 8 standards to the state laboratories until the 1990's it was not recommend to send in standards for periodic calibration through NIST. If an issue arose it was advised to clean the artifact in question to help to realign values in the control data (Harris, 2016).

To keep any large changes in mass from biasing the datasets all changes in mass were examined. The standard deviation of the individual dataset was determined and as a rule any changes in the data greater than two standard deviation would be flagged. No points in this study exceeded this limit and not data was dropped from the analysis.

The rate of usage of the standards is the final consideration. It is know that normal handling of mass standards causes ware and subsequent loss of mass. This is more evident when used on automated comparators where many more measurements are taken than if one was performing the same task by hand. Handling also presents the possibility of deposition of contaminates from the handling equipment.

5. Historic Shifts and Recalibrations Effecting Reported Mass values

Since the issuance of the Checkwate 8 kilograms to the States some corrections and recalibrations have been reported on the working standards of NBS/NIST, see Table 2.

On January 1, 1990 NIST announced a -0.170 mg/kg correction to be applied to all calibration performed prior to 1990. This was the result of the recalibration of K20 and K4 in 1985 and the concurrent calibration of set C (Davis, 1990).

On March 12, 2014 NIST published a significant addition in mass to K20 and K4 of +0.045 mg due to calibrations in 2010 and 2011 respectively. No advice was given at the time because it was not possible to determine when the shift had occurred (Kubarych & Abbott, 2014). This shift was soon to be counteracted by a more recent publication.

On March 24, 2015 it was published that the BIPM 'as-maintained' mass was offset from the IPK by 0.035 mg/kg (Stock, Barat, Davis, Picard, & Milton, 2015). On July 6, 2015 NIST Physical Measurement Laboratory sent out a letter advising customers to add a correction of -0.035 mg to all kilograms calibrated in between 2010 and 2015 and have them re-calibration in the near future (Kubarych, 2015).

Table 2: Reported Shifts

Date	Correction Value	Reason for Shift
1960 to 1990	-0.170 mg	Periodic recalibration of K20 and K4 in 1984
1990 to 2010		Nothing Reported
2010	0.045 mg	Periodic recalibration of K20 and K4 in 2010
2010 to 2015	-0.035 mg	BIMP Shift

6. Adjusting the Calibration Values for Reported Shifts in the Mass Scale

By accounting for the shifts mentioned in Section 3 and using the value as a correction for the NBS/NIST reported values, values were adjusted and there by reduced the noise seen in the data in Figure 1, see Figures 2. Again, the same method outlined in Section 2 was used to evaluate the stability of the artifacts and can be seen in Table 3.

Figure 2: Long Term Stability of Stainless Steel 1 kg State Standards With Initial Calibration Value Set to Zero and Adjusted for Reported Shifts in Mass

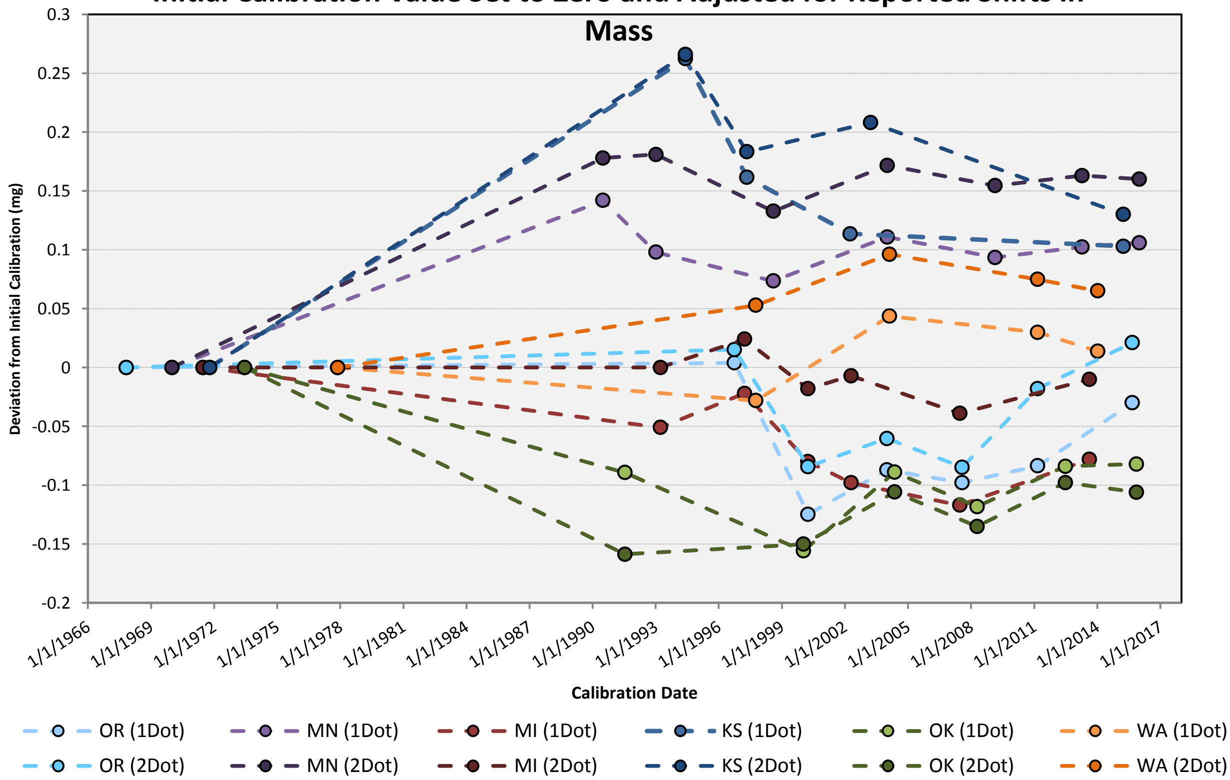

Table 3: Change in Mass per Year Adjusted for Reported Shifts in Mass

Data Set	Manufacturer and Issuance Year	Density (g/cm³)	Average Δm/year (mg)	Standard Deviation of Data (mg)
OR (1Dot) OR (2Dot)	W.&L.E. Gurley 1967	8.0017 8.0018	0.011 0.012	0.013 0.010
MN (1Dot) MN (2Dot)	Henry Troemner Inc. 1970	7.9926 7.9926	0.0061 0.0046	0.0055 0.0034
MI (1Dot) MI (2Dot)	Henry Troemner Inc. 1971	7.9925 7.9925	0.0071 0.0060	0.0033 0.0020
KS (1Dot) KS (2Dot)	Henry Troemner Inc. 1971	7.9921 7.9925	0.014 0.013	0.014 0.011
OK (1Dot) OK (2Dot)	Transmetric 1973	7.9920 7.9920	0.0074 0.0065	0.0048 0.0038
WA (1Dot) WA (2Dot)	NBS Instrument Shops 1977	7.9740 7.9740	0.0051 0.0040	0.0046 0.0019
Max Instability **Min Instability**	**Henry Troemner Inc.** **NBS Instrument Shops**	**7.9921** **7.9740**	**0.014** **0.0040**	**0.014** **0.0019**
Average Instability	**Checkwate 8**	**7.9740 to 8.0017**	**0.0080**	**0.0034**

The overall stability was determined by averaging the average change in mass per year of each artifact, the standard deviation was determined utilizing the same data.

7. Accounting for recalibrations of K20

United States kilogram prototype K20 has traveled to BIPM from time to time since its first issuance in 1889. Each calibration has produced a change in the true mass as cleaned. By accounting for each change in mass and combining as a cumulative change over time one can treat K20 as a 'stable' artifact by applying this change over time as a correction to each kilogram sets examined, see Table 4. This approach allows for another comparison to K20 and each kilogram pair can be examined relative to it, presented in Figure 3. K20 is kept in a highly controlled environment with its use being limited and is known to be more stable than the kilograms used on regular basis (Kubarych & Abbott, 2014).

Table 4: Cumulative Changes in True Mass of K20

Date	Change in Mass	Cumulative Change
1889	Issuance	0.000 mg
1937	+0.018 mg	+0.018 mg
1948	+0.002 mg	+0.020 mg
1984	-0.003 mg	+0.017 mg
1992	+0.001 mg	+0.018 mg
1999	-0.018 mg	0.000 mg
2010 *	+0.010 mg	+0.010 mg
2015	?	?

* 2010 values were adjusted for the -0.035 mg correction from BIPM

Figure 3: Long Term Stability of Stainless Steel 1 kg State Standards With Initial Calibration Value Set to Zero and Adjusting for Cumulative Shifts in K20

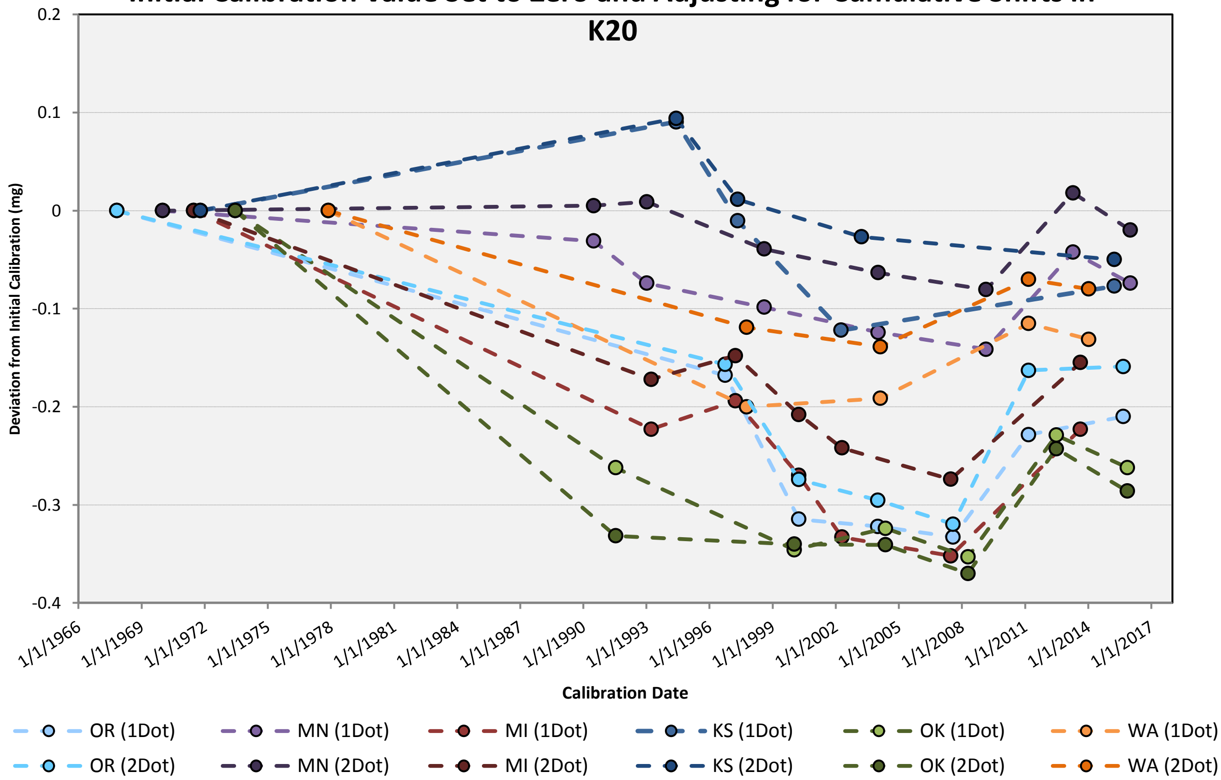

Table 5: Change in Mass per Year Adjusting for Cumulative Shifts in K20

Data Set	Manufacturer and Issuance Year	Density (g/cm³)	Average Δm/year (mg)	Standard Deviation of Data (mg)
OR (1Dot) OR (2Dot)	W.&L.E. Gurley 1967	8.0017 8.0018	0.017 0.018	0.014 0.016
MN (1Dot) MN (2Dot)	Henry Troemner Inc. 1970	7.9926 7.9926	0.0084 0.0084	0.0095 0.0080
MI (1Dot) MI (2Dot)	Henry Troemner Inc. 1971	7.9925 7.9925	0.015 0.006	0.017 0.013
KS (1Dot) KS (2Dot)	Henry Troemner Inc. 1971	7.9921 7.9925	0.015 0.012	0.016 0.010
OK (1Dot) OK (2Dot)	Transmetric 1973	7.9920 7.9920	0.009 0.011	0.013 0.012
WA (1Dot) WA (2Dot)	NBS Instrument Shops 1977	7.9740 7.9740	0.0044 0.0031	0.0070 0.0056
Max Instability **Min Instability**	**W.&L.E. Gurley** **NBS Instrument Shops**	**8.0018** **7.9740**	**0.018** **0.0031**	**0.016** **0.0056**
Average Instability	**Checkwate 8**	**7.9740 to 8.0017**	**0.0106**	**0.0049**

The overall stability was determined by averaging the average change in mass per year of each artifact, the standard deviation was determined utilizing the same data.

8. Conclusion

At first glance the data in Figure 1 appear highly unstable in comparison to the uncertainty each value carries. In fact, the kilograms are much more stable than the raw data lead one to believe and the results herein are similar to those stated in a previous study of ±0.005 mg/kg per year to ±0.020 mg/kg per year (Ueki, Mizishima, Sun, & Ueda, 2008). It was found that the Checkwate 8 mass standards have maintained an average stability of 0.0109 mg/kg per year with a standard deviation of 0.0036 mg as seen in Table 1. Even after accounting for the reported shifts in mass by NBS/NIST and treating K20 as a stable artifact the average instability changes minimally to 0.0080 mg/kg per year with a standard deviation of 0.0034 mg and to 0.0106 mg/kg per year with a standard deviation of 0.0049 mg respectively as presented in Tables 3 and 5 making the influences of the reported shifting of the mass scale through time difficult to filter from the raw data. Since the raw data and both methods of data manipulation provide values within their respected two standard deviation uncertainty bands it cannot be concluded what factor or combination of factors are the primary cause for the instability seen in the artifacts examined in this study. It should be further examined if laboratories working at the highest level of mass determination should start including a 0.010 mg/kg per year component of uncertainty to their overall uncertainty budget as the stability of artifact is not incorporated into the expanded uncertainties for mass calibrations provided by NIST.

The large shifts seen from 2010 to 2014 were due to the issues at BIPM and after following the advice issued by NIST the calibrations from that era fall in line with those before and after. The only way to further determine the stability of the Checkwate 8 mass standards is for more frequent calibrations of the artifacts themselves. A four-year calibration cycle is the current recommendation and has been adopted by most laboratories working at this level. This and the redefinition of the kilogram should provide more insight on this topic in future years.

9. Acknowledgments

Thanks you the laboratories who provided data for this study and the time spent digging up all of the original reports: Kevin Uphoff of the Kansas Metrology Laboratory, Nicholas Santini of the Michigan Department of Agriculture, Benjamin FitzPatrick of the Minnesota Standards Laboratory, Steve Sumner of the New Mexico Department of Agriculture, Richard Gonzales and Jeremy Nading of the Oklahoma Bureau of Standards, Aaron Aydelotte of the Oregon Department of Agriculture, Dan Wright of the Washington Department of Agriculture and finally Georgia Harris and Val Miller of National Institute of Standards and Technology for additional requests for information.

10. Works Cited

Allegheny Ludlum Steel Corporation. (1961). Checkwate-8. *158-42861*.

Davidson, S. (2003). A review of surface contamination and the stability of standard masses. *Metrologia*, 324-338.

Davidson, S. (2004). *The use of x-ray photoelectron spectroscopy to estimate the stability of primary mass standards.* NPL: Crown .

Davis, R. S. (1990). New Assignment of Mass Values and Uncertainties to NIST Working Standards. *J. Res. Natl. Inst. Stand. Technol. 95*.

Glaser, M., Borys, M., Ratschko, D., & Schwartz, R. (2010). Redfinition of the kilogram and the impact on its future dessemination. *Metrologia*, 419-428.

Grgic, G., Grun, M., Marti, K., & Berry, J. (2015). Report on application of the cleaning procedures to mass standards including gravimetric results. *EMRP SIB05 NewKILO WP4*.

Harris, G. (2016, January). Unpublished comunitcations.

Keysar, B. C. (1973, July 17). Letter to Dr. Earl Maxwell. Washington D.C.

Kubarych, Z. J. (2000). *Status of Mass Metrollogy at NIST in 2000.* Vienna, AU: XVI IMEKO World Congress.

Kubarych, Z. J. (2015, July 6). Letter notifying customers of shift due to calibrations at BIPM. Gaitherburg, Maryland: National Institute of Standards and Technology.

Kubarych, Z. J., & Abbott, P. J. (2014). The Dissemination of Mass in the United States: Results and Implications of Recent BIPM Calibrations of US National Prototype Kilograms. *J. Res. Natl. Inst. Stand. Technol. 119*.

Stabler, T. M. (1966). *State Standards and Laboratories.* Office of Weights and Measures, National Bureau of Standards.

Stock, M., Barat, P., Davis, R. S., Picard, A., & Milton, a. M. (2015). Calibration campaign against the international prototype of the kilogram in anticipation of the redefinition of the kilogram part 1: comparison of the internation prototype with its official copies. *Metrologia*, 310-316.

Ueki, M., Mizishima, S., Sun, J.-X., & Ueda, K. (2008). *Mass stability of 1-kg stainless steel weights in the past 13 years.* Japan: Univ. of Electro-Communications.